

us-09-525-041-2.rni

11/2/04

Alignment Scores:	
Pred. No.:	9,48e-18
Score:	218.00
Percent Similarity:	50.31%
Best Local Similarity:	32.92%
Query Match:	24.83%
	1
Length:	799
Matches:	53
Conservative:	28
Mismatches:	16
Indels:	64
Gaps:	7

Oy LeuLeuLeuSerGlyLeu---AlaIleuIleuValLeu---ProPhe 111
 Db 10 LeuLeuLeuSerGlyLeu---AlaIleuIleuValLeu---ProPhe 111
 Oy 79 ATGCTGCTTTCTGCTGCTCATGCTGCTGCTGCTGAGGTTCAAGGTGAAGAAACCCAGAGGAA 138
 Db 28 ---ProSer-----CysAlaProGlyIleTrpPheTyrIleIysSerAsnGlyTyr 42
 Oy 129 CTGCGCTCTGCACAGCATCGGCTGCTCCCAAGGCTCCAAAGGCTCAAGGCTATGAGCTCCACCTGCTAT 198
 Db 43 GLYTyrTrpPheArgIleLeuArgAsnTrpSerAspAlaGluLeuGluCysArgIleSerTyrGly 62
 Oy 199 GCCTTGATTTTTGTCACCAAAATCTCTGACAGATGAGATCTGGCTGGCAAGAGCGGCC 258
 Db 63 AsnGlyAlaIleLeuAlaSerIleLeuSerLeuIysGluValaSerThrIleAlaGluTyr 82
 Oy 259 TCTGGA---AACCTGGTGTCTGTGCTACAGGGGGCTGAAGGATCTTGCTGGTCCCTGCT 315
 Db 83 IleSerGlyTyrGlnArgSerGlnPro---IleTrpIleGlyLeuHisAspProGlnIys 101
 Oy 316 GTGAAGAGCATTTGGTAAACAGTACTACTACAGTCTGGATTTGGGCTTCATAGACCCACACAG 375
 Db 102 ArgIleGln-----TyrGlnTrpIleAspGlyAlaMetTyrLeuTyrArg 116
 Oy 376 GGCACCGAGCCCAATGAGAGAGATTGGAGAGTGGAGGTACAGCATGATGATGAATTAATTACTT 435
 Db 117 SerTrp-----SerGlyIysSerMetGlyIysAlaIleuValIysCysValaGluMetSerSer 134
 Oy 436 GCATGGGAGAGAAATCTCTCCACATCTCAAGCCCGGCGCACTGTGGAGAGCTGTCCAGA 495
 Db 135 AsnAsnAsnPheLeuThrTyrSerSerArgGlnCysValaIleuValSerGlnHisPheLeuCys 154
 Oy 496 AGCACAGCATTTCTGAGTGGAGAAAGATTATACGTAAATGTAGAGGTTACCTATGTCTGC 555
 Db 155 Lys 155
 Oy 556 AAA 558

RESULT 12
US-08-822-261-6
; Sequence 6, Application US/06822261
; Patent No. 5935813
; GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
CIVIL SERVICE K.

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36, 749
REFERENCE/DOCKET NUMBER: PF-0251 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 189600
;
US-08-822-261-6

Alignment Scores:
Pred. No.:      9,486-18
Score:          218.00
Best Similarity: 50.31%
Best Local Similarity: 32.92%
Query Match:    24.83%
DB:              2

length: 796
Matches: 53
Conservative: 28
Mismatches: 64
Indels: 16
Gaps: 7

```

[illegible]

L3 ANSWER 58 OF 59 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN
DUPLICATE 4
AN 1993:432762 BIOSIS
DN PREV199396087387
TI A gene homologous to the **reg** gene is
expressed in the human pancreas.
AU Bartoli, Catherine; Gharib, Bouchra; Giorgi, Dominique; Sansonetti, Alice;
Dagorn, Jean-Charles; Berge-LeFranc, Jean-Louis [Reprint author]
CS Unite 315 l'INSERM, 46 Boulevard Gaye, 13258 Marseille Cedex 9, France
SO FEBS (Federation of European Biochemical Societies) Letters, (1993) Vol.
327, No. 3, pp. 289-293.
CODEN: FEBLAL. ISSN: 0014-5793.
DT Article
LA English
OS Genbank-L08010
ED Entered STN: 22 Sep 1993
Last Updated on STN: 6 Nov 1993
AB We have determined the nucleotide sequence of **reg1** a human
genomic DNA fragment homologous to the **reg** gene which is
expressed in the exocrine pancreas and regenerating islets.
Sequence comparisons of **reg** and **reg1** suggested similar
exon-intron organisation. Based on this assumption, specific
oligonucleotides for **reg1** exons were used to demonstrate expression of the
reg1 gene in pancreas and liver, The proteins encoded by **reg** and
reg1 comprise 166 amino acids and differ by 22 amino acids only.

	CC	RA MEDLINE=94153997; PubMed=8110833; Morizumi S., Watanabe T., Uno M., Nakagawara K.I., Suzuki Y., Myashita H., Yonekura H., Okamoto H.; "Isolation, structural determination and expression of a novel reg gene, human regi beta."; Biochim. Biophys. Acta 1217:199-202(1994).
	RT	[2]
	RN	SEQUENCE FROM N.A.
	RP	MEDLINE=93351647; PubMed=8348956; Bertoli C., Gharib B., Giorgi D., Sansonetti A., Dagorn J.-C., Ra Berge-Lefranc U.; "A gene homologous to the reg gene is expressed in the human pancreas."; FEBS Lett. 327:289-293(1993).
	RL	[3]
	RP	CARBOHYDRATE-LINKAGE SITE
	RX	MEDLINE=95331286; PubMed=760722; De Reggi M., Capon C., Gharib B., Wieruszkeski J.M., Michel R., Fournet B.; "The glycan moiety of human pancreatic lithostathine. Structure characterization and possible pathophysiologic implications."; Eur. J. Biochem. 230:503-510(1995)
	CC	-1- FUNCTION: MIGHT ACT AS AN INHIBITOR OF SPONTANEOUS CALCIUM CARBONATE PRECIPITATION. MAY BE ASSOCIATED WITH NEURONAL SPROUTING IN BRAIN, AND WITH BRAIN AND PANCREAS REGENERATION.
	CC	-1- PWM: ALL O-LINKED GLYCANS CONSIST OF GAL-GLCNAc-GAL-GALNAc TETRASACCHARIDE CORE AND GET ELONGATED (MICROHETEROGENEITY).
	CC	-1- SIMILARITY: Contains 1 C-type lectin family domain.
	CC	-----
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
	CC	-----
	CC	EMBL; D17291; BAA04124.1; -
	CC	EMBL; D18816; BAA04091.1; -
	DR	EMBL; L08010; AAA18204.1; -
	DR	PIR; S34591; RGHLB.
	DR	HSP; P05451; ILIT.
	DR	Genew; HGNC:9952; REGIB.
	DR	MIM; 167771; -
	DR	GO; GO:008283; P:cell proliferation; TAS.
	DR	InterPro; IPR001304; Lectin_C.
	DR	InterPro; IPR003990; Pancreatic_ac.
	DR	Pfam; PF00059; lectin_c_1.
	DR	PRINTS; PR01504; PNCREXTITSAP.
	DR	SMART; SM00034; CLECT; 1.
	DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
	DR	PROSITE; PSS0041; C_TYPE_LECTIN_2; 1.
	KW	Glycoprotein; Signal; Lectin; Pyroglutamate carboxylic acid.
	FT	SIGNAL 1 22 BY SIMILARITY.
	FT	CHAIN 23 166 LITHOSTATHINE 1 BETA.
	FT	DOMAIN 34 164 C-TYPE LECTIN (LONG FORM).
	FT	MOD_RSS 23 23 PYROLIDONE CARBOXYLIC ACID (BY SIMILARITY).
	FT	CARBOHYD 27 27 O-LINKED (GALNAc...) (MUCIN TYPE).
	FT	DISULFID 36 47 BY SIMILARITY.
	FT	DISULFID 64 162 BY SIMILARITY.
	FT	DISULFID 137 154 BY SIMILARITY.
	QO	SEQUENCE 166 AA; 18665 MW; DIDC20E1IAESDBE8 CRC64;
	Query Match	27.3%; Score 240; DB 1; Length 166;
	Best Local Similarity	34.3%; Pred. No. 4,2e-17;
	Matches 46; Conservative % 27; Mismatches 57; Indels 4; Gaps 4;	

```

Matches 50; Conservative 33; Mismatches 72; Indels 19; Gaps 5
QY 1 MASSMALLLLSLCTATGVLDIT-----MRDSCAPGFYHKSCYCFRK 47
DB 1 MAQNNVYLILFLCMPLFSYSGQVAEDPFLAEKDLPSAKINCEGANAYSGCYLLIED 60
QY 48 LRNSDLELECOSYGNGAHJASILSLKEASTIAEYI--SGYORQPIWGLHDPQKROO 105
DB 61 RLTGAEADLPQNN-NAGHLVSLISQAEENFVASVSEGTASN-VWTGLHDPKSNRW 118
QY 106 QWIDAMNYLRWS--GKSMGNGKHCAMSSNNFLTWSSNECKRQHFLLCYR 157
DB 119 HWSGSLFLFKSWATGAPSTANRGYCVSLNSTNYKKWKDENCEAQSFWCKFR 172

RESULT 8
LECG_CROAT
ID _LECG_CROAT STANDARD; PRT; 135 AA.
AC P21963;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Galactose-specific lectin.
OS Crocalus atrox (western diamondback rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8730;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA MEDLINE=91115849; PubMed=1989986;
RA Hirabayashi J., Kusunoki T., Kasai K.-I.;
RT "Complete primary structure of a galactose-specific lectin from the
RT venom of the rattlesnake Crocalus atrox. Homologies with Ca2(+)-
RT dependent-type lectins."
RL J. Biol. Chem. 266:2320-2326(1991).
CC -1- FUNCTION: Galactose-binding protein which recognizes specific
CC carbohydrate structures and agglutinates a variety of animal cells
CC by binding to cell-surface glycoproteins and glycolipids. Calcium-
CC dependent lectin. Shows high hemagglutinating activity.
CC -1- SUBUNIT: Homodimer; disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSBP; P22897; IREG.
DR InterPro; IPR002353; AntifreezeZell.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatic_ac.
DR Pfam; PF00059; lectin_C; 1.
DR PRINTS; PR01504; PNCRENTISAP.
DR PRINTS; PR00356; ANTIFREEZEZIT.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PSS00645; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PSS00041; C_TYPE_LLECTIN_2; 1.
KW Calcium; Lectin.
FT DOMAIN 1 135 C-TYPE LECTIN (LONG FORM).
FT DISULFID 3 14
FT DISULFID 31 131
FT DISULFID 38 133
FT DISULFID 106 123
FT DISULFID 86
SQ SEQUENCE 135 AA; 16291 MW; 04BAC45DB2B721C8 CRC64; INTERCHAIN.

Query Match 26.5%; Score 233; DB 1; Length 135;
Best Local Similarity 34.4%; Prid. No. 1,7e-16;
Matches 45; Conservative 18; Mismatches 64; Indels 4; Gaps 2;
QY 29 SCAPGFYHKSNYGFRLKLRNSDLELECOSYGNGAHJASILSLKEASTIAEYISGYOR 88
DB 2 NCLDLWLPNNGGLCYKFLFNQDKTWEDAMFCRKYPQCHLASFPRYGESLEIAEYISDYRK 61
QY 89 SQE-1WGLHDPQKROOQWOMTIGAMNYLRWSGKS---NGKNGKHCAMSSNNFLTWSSN 144

```